# Lie series and invariant functions for analytic symplectic maps

Alex J. Dragt, and John M. Finn

Citation: Journal of Mathematical Physics **17**, 2215 (1976); doi: 10.1063/1.522868
View online: https://doi.org/10.1063/1.522868
View Table of Contents: http://aip.scitation.org/toc/jmp/17/12
Published by the American Institute of Physics

# Lie series and invariant functions for analytic symplectic maps*

Alex J. Dragt and John M. Finn†

*Center for Theoretical Physics, Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742*
(Received 3 March 1976)

Symplectic maps (canonical transformations) are treated from the Lie algebraic point of view using Lie series and Lie algebraic techniques. It is shown that under very general conditions an analytic symplectic map can be written as a product of Lie transformations. Under certain conditions this product of Lie transformations can be combined to form a single Lie transformation by means of the Campbell–Baker–Hausdorff theorem. This result leads to invariant functions and generalizes to several variables a classic result of Birkhoff for the case of two variables. It also provides a new approach since the connection between symplectic maps, Lie algebras, invariant functions, and Birkhoff's work has not been previously recognized and exploited. It is expected that the results obtained will be applicable to the normal form problem in Hamiltonian mechanics, the use of the Poincaré section map in stability analysis, and the behavior of magnetic field lines in a toroidal plasma device.

## 1. INTRODUCTION AND NOTATION

The purpose of this paper is to discuss canonical transformations from the Lie algebraic point of view using Lie series and Lie algebraic techniques. The study of canonical transformations or maps is important for several reasons: As is well known, canonical transformations preserve Hamilton's equations of motion.[1-3] In this context, they can be used systematically to bring a Hamiltonian to a simpler form from which the solutions to the equations of motion can be more easily discovered.[4-6] Secondly, the canonical coordinates $p(t)$, $q(t)$ at time $t$ for any Hamiltonian system are related to their values $p_0$, $q_0$ at time $t=t_0$ by a canonical transformation.[1-3] In addition, the Poincaré section map used to investigate stability behavior is canonical.[7,8] Finally, the behavior of magnetic field lines in a toroidal plasma device can be characterized by a canonical map.[9] We expect that our results will have application to all these areas.

The most commonly used method of describing canonical transformations is by means of transformation functions of mixed variables.[1-3] As has been discussed by Deprit and others, this method has certain drawbacks which can be overcome by the use of Lie series.[10-12] In this paper we will employ a variant of the Lie series approach.

The remainder of this section and Theorem 1 of the next section are devoted to a review of well-known material concerning Lie series and to a development of notation.[3,5,10-12] Our purpose is to make the material of this paper relatively self-contained.

We shall be working with a phase space consisting of the $2n$ variables $(q_1 \cdots q_n, p_1 \cdots p_n)$. The *Lie product* of any two functions $f$ and $g$ of the phase space variables will be defined by the Poisson bracket operation,

$$[f,g]=\sum_i \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}-\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}. \tag{1.1}$$

The set of all functions defined on phase space has an obvious linear vector space structure since it is closed under addition and scalar multiplication. Also the "multiplication" rule (1.1) satisfies all the requirements for a Lie product including the Jacobi condition

$$[f,[g,h]]+[g,[h,f]]+[h,[f,g]]=0. \tag{1.2}$$

Consequently, *functions on phase space may be viewed as elements in a Lie algebra.* We remind the reader that the equations of motion generated by a Hamiltonian $H$ can themselves be written in terms of Lie products,

$$\dot{q}_i=[q_i,H], \quad \dot{p}_i=[p_i,H]. \tag{1.3}$$

A canonical transformation to new variables $Q(q,p)$, $P(q,p)$ is defined to be any transformation satisfying

$$\begin{aligned}&[Q_i,Q_j]=[q_i,q_j]=0,\\&[P_i,P_j]=[p_i,p_j]=0,\\&[Q_i,P_j]=[q_i,p_j]=\delta_{ij}.\end{aligned} \tag{1.4}$$

That is, *canonical transformations are those transformations which preserve the Lie algebraic structure.*

At this point it is notationally convenient to collect the two sets of $n$ variables $q$, $p$ into a combined set of $2n$ variables $z_1 \cdots z_{2n}$ by the rule

$$z_i=q_i, \quad z_{n+i}=p_i, \quad i=1,\ldots,n. \tag{1.5}$$

In terms of the $z$'s the fundamental Poisson bracket rules (1.4) become

$$[z_i,z_j]=J_{ij}, \tag{1.6}$$

where $J$ denotes the antisymmetric $2n\times 2n$ matrix

$$J=\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}. \tag{1.7}$$

Here each entry in $J$ is an $n\times n$ block. We note that $J$ has the properties

$$\begin{aligned}&\tilde{J}=-J,\\&J^2=-I,\\&\det J=I.\end{aligned} \tag{1.8}$$

The general Lie product (1.1) is given in terms of the $z$'s by the relation

  

$$[f,g]=\sum_{k,l}\left(\frac{\partial f}{\partial z_k}\right) J_{kl}\left(\frac{\partial g}{\partial z_l}\right). \tag{1.9}$$

Suppose we introduce new variables $\bar{z}(z)$ and require that the transformation be canonical. Combining (1.4) and (1.9) we find

$$J_{ij}=[\bar{z}_i,\bar{z}_j]=\sum_{kl}\left(\frac{\partial \bar{z}_i}{\partial z_k}\right) J_{kl}\left(\frac{\partial \bar{z}_j}{\partial z_l}\right). \tag{1.10}$$

Let $M$ be the Jacobian matrix for the transformation going from $z$ to $\bar{z}$,

$$M_{ik}(z)=\frac{\partial \bar{z}_i}{\partial z_k}. \tag{1.11}$$

Employing $M$, we find that (1.10) can be written in the compact form

$$MJ\tilde{M}=J. \tag{1.12}$$

This is just the condition that the matrix $M$ must satisfy in order to belong to the symplectic group in $2n$ dimensions. We conclude that the necessary and sufficient condition for a transformation to be canonical is that its Jacobian matrix be symplectic.[8,12,13] For this reason a canonical transformation is often called a symplectic map.

In this paper we will study analytic symplectic maps. They are canonical transformations given by convergent power series. We write these power series as

$$\bar{z}_i=\sum_{|\sigma|>0} a_i(\sigma) z^\sigma. \tag{1.13}$$

Here $\sigma$ denotes a collection of exponents $(\sigma_1\cdots\sigma_{2n})$ and

$$|\sigma|=\sum_1^{2n}\sigma_i,\quad z^\sigma=z_1^{\sigma_1}z_2^{\sigma_2}\cdots z_{2n}^{\sigma_{2n}}. \tag{1.14}$$

Note that in the sum (1.13) we have purposely excluded constant terms by requiring $|\sigma|>0$. We do this to eliminate a possible nuisance later on and because we are not interested in transformations which simply translate the origin in phase space.

More specifically, our purpose is to study the relation between transformations of the form (1.13) and Lie series. Lie series and Lie transformations will be defined in the next section. There we will also see that under certain conditions the transformation (1.13) can be written as a product of Lie transformations. In Sec. 3 we will apply our results to several symplectic maps studied previously by other authors. Section 4 is devoted to the development and application of various Lie algebraic tools including the Campbell—Baker—Hausdorff formula. In Sec. 5 we apply the Campbell—Baker—Hausdorff formula to produce invariant functions for the map (1.13). An invariant function is a function $f$ with the property $f(\bar{z})=f(z)$. The existence and form of an invariant function enables one to study the effect of applying the map (1.13) many times in succession. We will learn that the determination of invariant functions is closely related to the determination of integrals of motion in Hamiltonian mechanics. In particular, invariant functions tell us a great deal about the underlying map just as integrals of motion characterize trajectories in mechanics. Our results are summarized in a final section.

## 2. LIE SERIES AND TRANSFORMATIONS

For the remainder of this paper we adopt the notational convention that lower case letters $f, g$, etc., denote functions and capital letters $F$, $G$, etc., denote operators.

Let $f$ be a specified function on phase space, and let $e$ be any function. We associate with $f$ the linear differential operator $F$ by the rule

$$Fe=[f,e]. \tag{2.1}$$

For example, if $f=z_1$, then $F=\partial/\partial z_{n+1}$. We shall call $F$ the *Lie operator* associated with $f$.

In general, Lie operators do not commute. Let $F$ and $G$ be the Lie operators associated with the functions $f$ and $g$. We will denote their commutator by $\{F, G\}$,

$$\{F,G\}=FG-GF. \tag{2.2}$$

Suppose $h$ is the function defined by

$$h=[f,g]. \tag{2.3}$$

We find, using the Jacobi relation (1.2),

$$\begin{aligned}\{F,G\}e&=[f,[g,e]]-[g,[f,e]]\\&=[[f,g],e]=He,\end{aligned} \tag{2.4}$$

where $H$ is the Lie operator associated with $h$. Since $e$ is any arbitrary function we may rewrite (2.4) as

$$H=\{F,G\}. \tag{2.5}$$

Comparing (2.5) and (2.3), we see that Lie operators form a Lie algebra under commutation which is homomorphic to the Poisson bracket Lie algebra of the underlying functions.[14] In particular, we are guaranteed that the commutator of two Lie operators is again a Lie operator. This fact will be important in Secs. 4 and 5.

We next consider infinite operator power series, called *Lie series*, of the form $\sum_0^\infty a_n F^n$ with the convention $F^0=I$. Of particular interest is the exponential series $\exp(F)$ defined as expected by

$$\exp(F)=\sum_0^\infty F^n/n!. \tag{2.6}$$

We shall call $\exp(F)$ the *Lie transformation associated* with $f$ and *generated* by $F$.

Lie transformations have two remarkable properties: Suppose $d$ and $e$ are any two functions. Then we find[10]

$$\exp(F)(de)=(\exp(F)d)(\exp(F)e) \tag{2.7}$$

and

$$\exp(F)[d,e]=[\exp(F)d,\exp(F)e]. \tag{2.8}$$

These results follow from the properties of the exponential series and the relations

$$F(de)=(Fd)e+d(Fe), \tag{2.9}$$

$$F[d,e]=[Fd,e]+[d,Fe]. \tag{2.10}$$

That is, $F$ is a *derivation* with respect to both ordinary and Poisson bracket multiplication.[15]

We are ready to explore the relation between symplectic maps and Lie transformations. The first result is immediate:

*Theorem* 1: If $\exp(F)$ is the Lie transformation associated with the analytic function $f$, then the infinite series given by

$$\bar{z}_i = \exp(F)\, z_i \tag{2.11}$$

is, providing it converges, an analytic symplectic map.

*Proof*: We simply use (2.8) and (2.6) to find

$$[\bar{z}_i, \bar{z}_j] = [\exp(F)\, z_i, \exp(F)\, z_j] = \exp(F)[z_i, z_j] = \exp(F)\, J_{ij} = J_{ij}. \tag{2.12}$$

The converse result is somewhat more difficult to state and to prove. We shall first state the result, and then work up to its proof in stages.

*Theorem* 2: Suppose one is given an analytic symplectic map in the form (1.13). Let $M(0)$ denote the matrix defined by (1.11) with all $z_i = 0$. Assume that $M(0)$ is joined to the identity matrix by a continuous one parameter subgroup of symplectic matrices. Or equivalently, assume that $M(0)$ can be written in the form

$$M(0) = \exp(JS), \tag{2.13}$$

where $S$ is a symmetric matrix. Then there exist homogeneous polynomials $g_2$, $g_3$, etc., of degree 2, 3, etc., and associated operators $G_2$, $G_3$, etc., such that the map (1.13) can be written in the infinite product form

$$\bar{z}_i = [\exp(G_2)\exp(G_3) \cdots]\, z_i. \tag{2.14}$$

The proof of this result is most easily accomplished by a series of lemmas.

*Lemma* 1: A set of $2n$ functions $f_1 \cdots f_{2n}$ satisfying

$$[z_i, f_j] = [z_j, f_i] \tag{2.15}$$

exists if and only if there is a function $g$ such that

$$f_i = [g, z_i] = G\, z_i. \tag{2.16}$$

*Proof*: First suppose that each $f_i$ is given by (2.16). Then we quickly verify (2.15): We find

$$[z_i, f_j] - [z_j, f_i] = [z_i, [g, z_j]] - [z_j, [g, z_i]] = -[g, [z_i, z_j]] = -[g, J_{ij}] = 0.$$

Now suppose (2.15) is true. We introduce auxiliary variables $z^*$ by the rule

$$z_i^* = \sum_j J_{ij} z_j. \tag{2.17}$$

Because $\tilde{J}J = I$, we can immediately write the inverse relation

$$z_k = \sum_i J_{ik} z_i^*. \tag{2.18}$$

Let $f$ be any function. We find

$$[z_i, f] = \sum_{jk} \left(\frac{\partial z_i}{\partial z_j}\right) J_{jk} \left(\frac{\partial f}{\partial z_k}\right) = \sum_k J_{ik} \left(\frac{\partial f}{\partial z_k}\right) = \sum_k \left(\frac{\partial f}{\partial z_k}\right)\left(\frac{\partial z_k}{\partial z_i^*}\right) = \left(\frac{\partial f}{\partial z_i^*}\right). \tag{2.19}$$

Because of this relation, the hypothesis (2.15) implies

$$\frac{\partial f_j}{\partial z_i^*} = \frac{\partial f_i}{\partial z_j^*}$$

which means that $\sum_i f_i\, dz_i^*$ is an exact differential. Therefore the function $g$ given by the path integral

$$g = -\int^{z^*} \sum_i f_i\, dz_i'^* \tag{2.20}$$

is well defined, and satisfies $[g, z_i] = -(\partial g/\partial z_i^*) = f_i$. Using (2.17) we obtain the explicit formula

$$g(z) = -\int^{z} \sum_{ij} f_i J_{ij}\, dz_j'. \tag{2.21}$$

*Lemma* 2: Let $g_s$ be a homogeneous polynomial of degree $s$. That is, we have

$$g_s(z) = \sum_{|\sigma|=s} b(\sigma)\, z^\sigma \tag{2.22}$$

for some set of coefficients. Also, let $\mathcal{P}_s$ denote the set of all homogeneous polynomials of degree $s$. Then, since the Poisson bracket operation involves multiplication and two differentiations, we have for any two homogeneous polynomials $g_r$, $g_s$ the relation

$$[g_r, g_s] \in \mathcal{P}_{r+s-2}. \tag{2.23}$$

*Lemma* 3: A necessary and sufficient condition for a symplectic matrix $N$ to lie on a continuous one parameter symplectic subgroup joined to the identity is that there exist a symmetric matrix $S$ such that

$$N = \exp(JS). \tag{2.24}$$

*Proof*: Suppose $N$ is a matrix of the form (2.24). Then we find by direct computation that $N$ is symplectic,

$$N J \tilde{N} = \exp(JS)\, J \exp(JS)^{\sim} = \exp(JS)\, J \exp(S\tilde{J}) = \exp(JS)\, J \exp(-SJ)\, J^{-1} J = \exp(JS)\exp(-JS)\, J = J. \tag{2.25}$$

A similar result holds for the matrix $N(\tau)$ defined by

$$N(\tau) = \exp(\tau JS) \tag{2.26}$$

where $\tau$ is a parameter. It follows that $N$ lies on a continuous one parameter subgroup joined to the identity. Now assume the converse, namely that $N$ does lie on a one parameter subgroup. Without loss of generality we assume that the group is parameterized in such a way that

$$N(\tau_1 + \tau_2) = N(\tau_1)\, N(\tau_2), \tag{2.27a}$$

$$N(0) = I, \tag{2.27b}$$

$$N(1) = N. \tag{2.27c}$$

Now differentiate (2.27a) with respect to $\tau_1$ and then set $\tau_1 = 0$ and $\tau_2 = \tau$ to obtain the result

$$N'(\tau) = N'(0)\, N(\tau). \tag{2.28a}$$

This equation with the initial condition (2.27b) has the unique solution

$$N(\tau) = \exp[\tau N'(0)]. \tag{2.28b}$$

Let us write $N'(0) = JS$ where $S$ is an undetermined matrix. Next suppose $\tau$ is small. Then

$$N(\tau) = \exp(\tau JS) \cong I + \tau JS. \tag{2.29a}$$

Enforcing the symplectic condition (1.12) gives

$$(I + \tau JS)\, J (I + \tau \tilde{S}\tilde{J}) \cong J. \tag{2.29b}$$

Consequently, equating powers of $\tau$, we have

$$JSJ + \tilde{J}\tilde{S}\tilde{J} = 0. \tag{2.29c}$$

Finally, use of (1.8) implies the expected conclusion

$$S = \tilde{S}. \tag{2.30}$$

Now set $\tau = 1$. The result is a matrix written in the form (2.24).

*Cautionary remark*: Not every symplectic matrix can be written in the form (2.24). A counter example in the $2\times 2$ case is the matrix given by

$$N = \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix}. \tag{2.31}$$

*Lemma* 4: Suppose that $M(0)$ is joined to the identity by a continuous one parameter subgroup. Then there exists a second degree homogeneous polynomial $g_2$ such that

$$(\exp G_2)\, z_i = \sum_j M_{ij}(0)\, z_j. \tag{2.32}$$

*Proof*: According to Lemma 3 we may write $M(0)$ in the form

$$M(0) = \exp(JS). \tag{2.33}$$

We define $g_2$ by the expression

$$g_2 = -\tfrac{1}{2} \sum_{i,k} S_{ik} z_i z_k, \tag{2.34}$$

and find

$$G_2 z_i = [g_2, z_i] = \sum_j (JS)_{ij} z_j. \tag{2.35}$$

The desired result (2.32) follows immediately by exponentiation.

*Lemma* 5: Let $r(>1)$ denote a "remainder" series consisting of terms higher than first degree. Then, under the conditions of Theorem 2,

$$\exp(-G_2)\, \bar{z}_i = z_i + r(>1). \tag{2.36}$$

*Proof*: From (1.13) we have

$$\exp(-G_2)\, \bar{z}_i = \sum_{|\sigma|=1} a_i(\sigma) \exp(-G_2)\, z^\sigma + r(>1). \tag{2.37}$$

Since the first term on the right is of first order, we can also equivalently write

$$\exp(-G_2)\, \bar{z}_i = \sum_j M_{ij}(0) \exp(-G_2)\, z_j + r(>1), \tag{2.38}$$

but from (2.32) we conclude

$$\exp(-G_2)\, z_j = \sum_k (M^{-1})_{jk} z_k. \tag{2.39}$$

Combining (2.38) and (2.39) completes the proof.

*Lemma* 6: There exist polynomials $g_3$, $g_4$, etc., such that when $\exp(-G_3)$, $\exp(-G_4)$, etc., are consecutively applied to (2.36), the order of the remainder term can be made arbitrarily large.

*Proof*: We shall find $g_3$. The higher order $g$'s are found in the same fashion. Let us decompose the remainder term $r(>1)$ in (2.36) into a second degree term $f_i(2;z)$ plus a higher order remainder $r(>2)$,

$$\exp(-G_2)\, \bar{z}_i = z_i + f_i(2;z) + r(>2). \tag{2.40}$$

Now form the Poisson bracket of (2.40) with the analogous expression having the index set equal to $j$. Using (1.10), (2.8), and (2.23) we find

$$J_{ij} = [z_i + f_i(2) + r(>2),\; z_j + f_j(2) + r(>2)],$$

or

$$J_{ij} = J_{ij} + [z_i, f_j(2)] + [f_i(2), z_j] + r(>1). \tag{2.41}$$

Equating like powers of $z$, we get

$$[z_i, f_j(2)] + [f_i(2), z_j] = 0. \tag{2.42}$$

It follows from Lemma 1 that there is a function $g_3$ such that

$$f_i(2;z) = G_3 z_i. \tag{2.43}$$

In fact, $g_3$ can be found explicitly from (2.21), and is clearly homogeneous of degree 3. Using (2.43), we rewrite (2.40) as

$$\exp(-G_2)\, \bar{z}_i = z_i + G_3 z_i + r(>2). \tag{2.44}$$

Finally, we apply $\exp(-G_3)$ to both sides of (2.44). The result is

$$\exp(-G_3)\exp(-G_2)\, \bar{z}_i = z_i + r(>2). \tag{2.45}$$

We have all the necessary machinery to complete the proof of Theorem 2. Comparing (2.36) and (2.45), we find that we have been able to raise the order of the remainder term by 1. As stated in the last lemma, it is easy to see that the process can be repeated at will. That is, there exist further homogeneous polynomials $g_4, g_5, \ldots, g_s$ such that

$$\exp(-G_s)\cdots\exp(-G_3)\exp(-G_2)\, \bar{z}_i = z_i + r(>s-1) \tag{2.46}$$

for any value of $s$. Inverting the left-hand side of (2.46), we obtain the result

$$\bar{z}_i = \exp(G_2)\cdots\exp(G_s)\, z_i + r(>s-1). \tag{2.47}$$

Now let $s \to \infty$. Then, if the remainder tends to zero, we obtain the advertised result (2.14). Otherwise the result is true only formally. In the latter case the infinite product is also divergent.

We close this section with the remark that it is often more convenient to have a product representation, usually with different $G$'s, in the opposite order,

$$\bar{z}_i = \exp(G_s'')\cdots\exp(G_3'')\exp(G_2)\, z_i + r(>s-1). \tag{2.48}$$

We will show in Sec. 4 that this is always possible providing (2.47) holds, and vice versa.

## 3. EXAMPLES

In this section we will apply the results of Sec. 2 to some maps studied previously in the literature by other authors.

The first examples are *Cremona* maps. They are symplectic maps for which the power series (1.13) terminates.[16] The simplest nontrivial Cremona map terminates at the second power. In the easiest case of two dimensions, where symplectic maps are merely area preserving maps, a suitable linear transformation brings the quadratic Cremona map into the form[17]

$$\bar{z}_1 = \lambda[z_1 + (z_1 - z_2)^2], \quad \bar{z}_2 = \lambda^{-1}[z_2 + (z_1 - z_2)^2] \tag{3.1}$$

if the matrix $M(0)$ for the original map has the real distinct positive eigenvalues $\lambda$ and $\lambda^{-1}$. (Note that for a symplectic matrix $M$, one always has $\det M = 1$, and hence the eigenvalues must be reciprocals in the $2\times 2$ case.) If the eigenvalues are $\exp(\pm i\alpha)$, i.e., if they lie on the unit circle, then the quadratic Cremona map can be brought to the form[17–19]

$$\begin{aligned} \bar{z}_1 &= z_1 \cos\alpha + z_2 \sin\alpha + z_2^2 \cos\alpha, \\ \bar{z}_2 &= -z_1 \sin\alpha + z_2 \cos\alpha - z_2^2 \sin\alpha. \end{aligned} \tag{3.2}$$

There are also other possibilities for the eigenvalues of $M(0)$ which are of less interest.

Let us apply our formalism. In the case (3.1) we have

$$M(0) = \begin{pmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{pmatrix} = \exp\begin{pmatrix} \log\lambda & 0 \\ 0 & -\log\lambda \end{pmatrix}. \tag{3.3}$$

Therefore, using (2.33) and (2.34) we find

$$g_2 = -(\log\lambda)\, z_1 z_2. \tag{3.4}$$

Correspondingly, we have for $G_2$ the expression

$$G_2 = (\log\lambda)\left(z_1 \frac{\partial}{\partial z_1} - z_2 \frac{\partial}{\partial z_2}\right). \tag{3.5}$$

Next we compute the $f_i(2;z)$ following (2.40). We find

$$\begin{aligned} \exp(-G_2)\,\bar{z}_1 &= \exp(-G_2)\,\{\lambda[z_1 + (z_1 - z_2)^2]\} \\ &= \lambda \exp(-G_2)\, z_1 + \lambda[\exp(-G_2)\, z_1 - \exp(-G_2)\, z_2]^2 \\ &= z_1 + \lambda[\lambda^{-1} z_1 - \lambda z_2]^2. \end{aligned} \tag{3.6}$$

Thus, we get for $f_1(2;z)$ the expression

$$f_1(2;z) = \lambda^{-1} z_1^2 - 2\lambda\, z_1 z_2 + \lambda^3 z_2^2. \tag{3.7}$$

Similarly, we find for $f_2(2;z)$ the result

$$f_2(2;z) = \lambda^{-3} z_1^2 - 2\lambda^{-1} z_1 z_2 + \lambda z_2^2. \tag{3.8}$$

We are ready to apply (2.21) to find $g_3$. The line integral is most easily evaluated along the path $z_i' = \tau z_i$ with the parameter $\tau$ ranging from zero to one. If the $f_i$ are homogeneous polynomials, the integral can be evaluated immediately in the general case to give

$$g_{s+1}(z) = -(s+1)^{-1} \sum_{i,j} f_i(s;z) J_{ij} z_j. \tag{3.9}$$

In particular, for $g_3$ we find the result

$$g_3(z) = (\lambda^{-1} z_1 - \lambda z_2)^3/3. \tag{3.10}$$

It follows that

$$G_3 = (\lambda^{-1} z_1 - \lambda z_2)^2 \left(\lambda^{-1} \frac{\partial}{\partial z_2} + \lambda \frac{\partial}{\partial z_1}\right). \tag{3.11}$$

We must now continue on to compute the higher order remainder terms following (2.45). The calculation is simplified by the observation that

$$G_3^2 z_i = 0 \tag{3.12}$$

and hence

$$\exp(-G_3)\, z_1 = z_1 - \lambda(\lambda^{-1} z_1 - \lambda z_2)^2, \tag{3.13a}$$

$$\exp(-G_3)\, z_2 = z_2 - \lambda^{-1}(\lambda^{-1} z_1 - \lambda z_2)^2. \tag{3.13b}$$

Also, we have

$$G_3(\lambda^{-1} z_1 - \lambda z_2)^2 = 0 \tag{3.14}$$

and hence

$$\exp(-G_3)(\lambda^{-1} z_1 - \lambda z_2)^2 = (\lambda^{-1} z_1 - \lambda z_2)^2. \tag{3.15}$$

We are ready. We find, using (3.6), (3.13), and (3.15) the result

$$\exp(-G_3)\exp(-G_2)\,\bar{z}_1 = \exp(-G_3)[z_1 + \lambda(\lambda^{-1} z_1 - \lambda z_2)^2] = z_1. \tag{3.16}$$

That is, the remainder term vanishes! The same is true for $\bar{z}_2$. Thus in this case, the higher order Lie operators $G_4$, $G_5$, etc., are all zero. We conclude that for the two-dimensional quadratic Cremona map in the case (3.1) we have

$$\exp(-G_3)\exp(-G_2)\,\bar{z}_i = z_i, \tag{3.17}$$

and hence

$$\bar{z}_i = \exp(G_2)\exp(G_3)\, z_i. \tag{3.18}$$

The calculation in the case (3.2) can also be carried out with equal ease. The result is

$$g_2(z) = -(\alpha/2)(z_1^2 + z_2^2), \tag{3.19}$$

$$G_2 = -\alpha\left(z_1 \frac{\partial}{\partial z_2} - z_2 \frac{\partial}{\partial z_1}\right), \tag{3.20}$$

$$g_3(z) = -(z_1 \sin\alpha + z_2 \cos\alpha)^3/3, \tag{3.21}$$

$$G_3 = -(z_1 \sin\alpha + z_2 \cos\alpha)^2 \left(\sin\alpha \frac{\partial}{\partial z_2} - \cos\alpha \frac{\partial}{\partial z_1}\right). \tag{3.22}$$

The higher order Lie operators again vanish, and Eq. (3.18) is exact.

Another symplectic map which has received considerable study is the ninth order Cremona map in two dimensions given implicitly by the relations[20,21]

$$\bar{z}_1 = z_1 + a z_2 - a z_2^3, \quad \bar{z}_2 = z_2 - a\bar{z}_1 + a\bar{z}_1^3 \tag{3.23}$$

and explicitly by

$$\begin{aligned} \bar{z}_1 &= z_1 + a z_2 - a z_2^3, \\ \bar{z}_2 &= z_2 - a(z_1 + a z_2 - a z_2^3) + a(z_1 + a z_2 - a z_2^3)^3. \end{aligned} \tag{3.24}$$

Here $a$ is a parameter. Due to algebraic complications, we have not attempted to express this map in the form (2.14) although we have verified that $M(0)$ does lie on a continuous one parameter subgroup connected to the identity providing $a$ is small enough. This task seems better suited to digital computers programmed to perform algebraic operations. However, we have discovered that the map can be written in the form

$$\bar{z}_i = \exp(F_2)\exp(F_4)\exp(G_2)\exp(G_4)\, z_i \tag{3.25}$$

with

$$\begin{aligned} f_2 &= -(a/2)\, z_2^2, \quad f_4 = (a/4)\, z_2^4, \\ g_2 &= -(a/2)\, z_1^2, \quad g_4 = (a/4)\, z_1^4. \end{aligned} \tag{3.26}$$

We will see in the next section that, should it be desirable, there are standard Lie algebraic manipulations which can be used to bring (3.25) into the form (2.14).

As mentioned by Moser,[22] two-dimensional Cremona mappings can be expressed as repeated products of linear transformations and shear mappings of the form

$$\bar{z}_1 = z_1 + h(z_2), \quad \bar{z}_2 = z_2. \tag{3.27}$$

The shear mapping (3.27) can be expressed as a Lie series $\exp(F)$, where $F$ is the Lie operator associated with the function $f$ given by

$$f(z_1, z_2) = -\int^{z_2} h(z')\,dz'. \tag{3.28}$$

Also, linear maps connected to the identity can be expressed as Lie transformations as seen earlier. It follows that quite generally Cremona maps can be expressed as products of Lie series. This factorization may be distinct from that of (2.14) since the functions given by (3.28) need not be homogeneous polynomials.

In addition, some area preserving maps which are more general than Cremona maps can be factored in a similar way. For example, the mapping $\mathcal{T}$ given by[23]

$$\mathcal{T}:\begin{cases}\bar{z}_1 = z_1 + z_2,\\ \bar{z}_2 = z_1 - \epsilon\sin(z_1+z_2)\end{cases} \tag{3.29}$$

can be expressed as a product of two shear mappings. We have

$$\mathcal{T} = \mathcal{R}\mathcal{S}, \tag{3.30}$$

where $\mathcal{R}$ and $\mathcal{S}$ denote the mappings

$$\mathcal{R}:\begin{cases}\bar{z}_1 = z_1,\\ \bar{z}_2 = \bar{z}_2 - \epsilon\sin z_1,\end{cases} \tag{3.31}$$

$$\mathcal{S}:\begin{cases}\bar{z}_1 = z_1 + z_2,\\ \bar{z}_2 = z_2.\end{cases} \tag{3.32}$$

The mapping $\mathcal{T}$ therefore has the representation

$$\bar{z}_i = \exp(F)\exp(G)\,z_i, \tag{3.33}$$

where $F$ and $G$ are the Lie operators corresponding to the functions $f$ and $g$ given by

$$f(z_1, z_2) = \epsilon\cos z_1, \quad g(z_1, z_2) = -z_2^2/2. \tag{3.34}$$

Similar results hold for the mapping

$$\bar{z}_1 = z_1 + \epsilon\sin z_2, \quad \bar{z}_2 = z_1 + z_2 + \epsilon\sin z_2 \tag{3.35}$$

studied by Froeschlé.[18] It can be written in the product form $\mathcal{S}\mathcal{R}$ provided the roles of $z_1$ and $z_2$ are interchanged.

Finally, we close this section with a brief study of the four-dimensional map given by[24]

$$\begin{aligned}\bar{z}_1 &= z_1 + a_1\sin(z_1+z_3) + b\sin(z_1+z_2+z_3+z_4),\\ \bar{z}_2 &= z_2 + a_2\sin(z_2+z_4) + b\sin(z_1+z_2+z_3+z_4),\\ \bar{z}_3 &= z_1 + z_3,\\ \bar{z}_4 &= z_2 + z_4.\end{aligned} \tag{3.36}$$

A routine calculation shows that $M(z)$ given by (1.11) is indeed a symplectic matrix so that (3.36) is a symplectic map. In particular, $M(0)$ is given by the matrix

$$M(0) = \begin{pmatrix} (1+a_1+b) & b & (a_1+b) & b\\ b & (1+a_2+b) & b & (a_2+b)\\ 1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\end{pmatrix}. \tag{3.37}$$

We next inquire whether $M(0)$ lies on a continuous one parameter subgroup connected to the identity. We have not been able to treat the general case. However, we have been able to verify this condition if $a_1$, $a_2$, and $b$ are sufficiently small. Details are given in Appendix A. Therefore Theorem 2 applies, and with sufficient effort the polynomials $g_2$, $g_3$, etc., can be computed.

## 4. LIE ALGEBRAIC TOOLS

The content of Theorem 2 is that under rather general conditions a symplectic map can be written as a product of Lie transformations. Earlier, from Eqs. (2.2)–(2.5), we found that the Lie operators which generate Lie transformations also form a Lie algebra under commutation. The purpose of this section is to review some Lie algebraic tools which will enable us to manipulate the various Lie operators appearing in products of Lie transformations.

We begin by introducing yet another Lie algebra, the *adjoint* Lie algebra.[14] Let $F$ be a given Lie operator and $E$ an arbitrary Lie operator. We associate with $F$ an operator $\hat{F}$ (which acts on Lie operators) by the rule

$$\hat{F}E = \{F, E\}. \tag{4.1}$$

Here, as before, $\{\ ,\ \}$ denotes commutation. Next, let $F$ and $G$ be any two Lie operators. We define a Lie operator $H$ by the rule

$$H = \{F, G\}. \tag{4.2}$$

Then we find

$$\begin{aligned}\{\hat{F}, \hat{G}\}E &= (\hat{F}\hat{G} - \hat{G}\hat{F})E\\ &= \{F, \{G, E\}\} - \{G, \{F, E\}\}\\ &= \{\{F, G\}, E\} = \hat{H}E.\end{aligned} \tag{4.3}$$

Here we have used the Jacobi identity

$$\{E, \{F, G\}\} + \{F, \{G, E\}\} + \{G, \{E, F\}\} = 0 \tag{4.4}$$

which always holds for commutators. Since $E$ is arbitrary, we may rewrite (4.3) as

$$\hat{H} = \{\hat{F}, \hat{G}\}. \tag{4.5}$$

Comparison of (4.5) and (4.2) shows that the adjoint Lie algebra is homomorphic to the parent Lie algebra of Lie operators.

Our discussion should have a familiar ring. It parallels, in fact, the discussion surrounding Eqs. (2.1)–(2.5). Reviewing these equations, we see that the commutator Lie algebra of Lie operators is actually the adjoint Lie algebra of the underlying Poisson bracket Lie algebra. And, consequently, the "adjoint" we have been discussing is really the "adjoint-adjoint" of the basic Poisson bracket Lie algebra.

We now have the necessary notation to state the simplest theorem about the rearrangement of Lie transformations,

*Theorem* 3: Let $A$ and $B$ be Lie operators. Then

$$[\exp(A)]B[\exp(-A)]=(\exp\hat{A})B \tag{4.6}$$

and

$$\exp(A)\exp(B)\exp(-A)=\exp[(\exp\hat{A})B]. \tag{4.7}$$

*Proof*: Let $\tau$ be a parameter, and define $C(\tau)$ by the equation

$$C(\tau)=[\exp(\tau A)]B[\exp(-\tau A)]. \tag{4.8}$$

Then we have the relation

$$C(0)=B. \tag{4.9}$$

Further, we find by differentiation

$$\frac{dC}{d\tau}=AC-CA=\hat{A}C. \tag{4.10}$$

The solution to this differential equation with the boundary condition (4.9) is

$$C(\tau)=\exp(\tau\hat{A})B. \tag{4.11}$$

Now set $\tau=1$ to obtain (4.6). We next observe that for any two operators $B$ and $C$ we have

$$\hat{A}(BC)=\{A,BC\}=\{A,B\}C+B\{A,C\}=(\hat{A}B)C+B\hat{A}C. \tag{4.12}$$

It follows that $\hat{A}$ acts as a derivation on products, and in analogy to (2.7) and (2.9) we obtain

$$(\exp\hat{A})(BC)=((\exp\hat{A})B)((\exp\hat{A})C). \tag{4.13}$$

This result is consistent with (4.6) and the observation that

$$[\exp(A)](BC)[\exp(-A)]$$
$$=[\exp(A)]B[\exp(-A)][\exp(A)]C[\exp(-A)]. \tag{4.14}$$

We conclude that

$$[\exp(A)]B^n[\exp(-A)]=[(\exp\hat{A})B]^n \tag{4.15}$$

for any power $n$. The desired result (4.7) now follows directly term by term.

As an application of Theorem 3, let us consider the product $\exp(G_2)\exp(G_3)$ which occurs in Eq. (3.18). We have

$$\begin{aligned}\exp(G_2)\exp(G_3)&=\exp(G_2)\exp(G_3)\exp(-G_2)\exp(G_2)\\&=\exp[(\exp\hat{G}_2)G_3]\exp(G_2).\end{aligned} \tag{4.16}$$

Let us define $G_3'$ by the expression

$$G_3'=(\exp\hat{G}_2)G_3. \tag{4.17}$$

Is there a polynomial $g_3'$ which has $G_3'$ as its associated Lie operator? We know there must be from the remarks following Eq. (2.5). Using the homomorphisms between the Lie algebras involved, we obtain

$$g_3'=(\exp G_2)g_3. \tag{4.18}$$

Consequently, for the first example of Sec. 3 we calculate that

$$g_3'=(z_1-z_2)^3/3 \tag{4.19}$$

and hence

$$G_3'=(z_1-z_2)^2\left(\frac{\partial}{\partial z_2}+\frac{\partial}{\partial z_1}\right). \tag{4.20}$$

We have shown that

$$\exp(G_2)\exp(G_3)=\exp(G_3')\exp(G_2) \tag{4.21}$$

with $G_3'$ given by (4.20).

We are now ready to move on to a far deeper result generally known as the Campbell—Baker—Hausdorff (CBH) formula. In its usual mathematical setting it provides the connecting link between Lie algebras and Lie groups.[14] We will use it to reexpress the product of two Lie transformations as a single Lie transformation, or more generally as a method of combining exponents.

*Theorem* 4: Let $A$ and $B$ be any two operators, and let $\alpha$ and $\beta$ be parameters. Then we can formally write

$$\exp(\alpha A)\exp(\beta B)=\exp(C) \tag{4.22}$$

with

$$\begin{aligned}C=\alpha A+\beta B+(\alpha\beta/2)\{A,B\}+(\alpha^2\beta/12)\{A,\{A,B\}\}\\+(\alpha\beta^2/12)\{B,\{B,A\}\}+\cdots.\end{aligned} \tag{4.23}$$

The series for $C$ may or may not converge depending on the properties of $\alpha A$ and $\beta B$. The really remarkable fact is that the right-hand side of (4.23) involves *only* Lie products. Thus, all we need to evaluate (4.23) is a knowledge of the Lie algebra generated by $A$ and $B$, and we are guaranteed that $C$ is contained in this Lie algebra. The general form of all the coefficients in the series is not known.[25] However, the series can be formally summed to all orders in $\alpha$ and the first few orders in $\beta$. Through first order in $\beta$ we have

$$C=\alpha A+\beta\alpha\hat{A}[1-\exp(-\alpha\hat{A})]^{-1}B+O(\beta^2). \tag{4.24}$$

A proof of these results and an expression for the quadratic term in $\beta$ are given in Appendix B.

As a simple example of the use of the CBH formula, we will derive (2.48) starting from (2.47). A more complicated example of its use will be given in the next section. Beginning with (2.47), we write

$$\begin{aligned}\exp(G_2)\cdots\exp(G_s)&=\exp(G_2)\cdots\exp(G_s)\exp(-G_2)\exp(G_2)\\&=\exp(G_3')\cdots\exp(G_s')\exp(G_2),\end{aligned} \tag{4.25}$$

where

$$G_r'=\exp(\hat{G}_2)G_r. \tag{4.26}$$

Note that as in our earlier example, $G_r'$ will be the Lie operator associated with the function $g_r'$ given by

$$g_r'=(\exp G_2)g_r, \tag{4.27}$$

and that the degree of $g_r'$ is as indicated because of (2.23). Next we repeatedly use the CBH formula (4.23) to combine the various operators $G_r'$ to obtain an expression of the form

$$\exp(G_3')\cdots\exp(G_s')=\exp(H_3+\cdots+H_s+\cdots). \tag{4.28}$$

Observe that because of (2.23), only a finite number of terms in the series (4.23) are required in the calculation of each $H_r$. Again using the CBH formula, we may write

$$\exp(H_3+\cdots)=\exp(H_3+\cdots)\exp(-H_3)\exp(H_3)$$
$$=\exp(H_4'+\cdots)\exp(H_3). \quad (4.29)$$

This process can be repeated again and again to get

$$\exp(H_3+\cdots)=\exp(H_{s+1}''+\cdots)\exp(H_s')\cdots\exp(H_3). \quad (4.30)$$

Combining (4.25), (4.28), and (4.29), we find

$$\exp(G_2)\cdots\exp(G_s)z_i$$
$$=\exp(H_s')\cdots\exp(H_3)\exp(G_2)z_i+r(>s-1). \quad (4.31)$$

Consequently, an expression of the form (2.47) implies an expression of the form (2.48). The converse can be proven analogously.

## 5. CONSTRUCTION OF INVARIANT FUNCTIONS

In the study of a symplectic transformation $\mathcal{T}$ arising from either a Poincaré surface of section or from following the field lines in a toroidal plasma device, one is interested in studying the result of applying the map many times in succession. That is, we are interested in studying $\mathcal{T}^n z$ for large $n$. This study is simplified if one can construct invariant functions $f$ with the property

$$f(\bar{z})=f(z), \quad (5.1)$$

where

$$\bar{z}=\mathcal{T}z. \quad (5.2)$$

For if such functions can be found, one knows that the points generated by $\mathcal{T}^n z$, for various $n$ must all lie on a surface of constant $f$. The more invariant functions one can find, the more one can say about the map and its powers. The situation is quite analogous to the role played by integrals of motion for Hamiltonian systems or magnetic surfaces in a toroidal plasma. We will see shortly that the analogy is more than coincidental.

The problem of constructing invariant functions in the case of symplectic transformations in *two* variables was first considered in detail by Birkhoff.[26] We shall begin this section by proving some simple lemmas which will enable us to restate his result in our language. We will then show how the same results can be obtained *for any number of variables* from the CBH formula.

*Lemma* 7: Consider a one parameter family of symplectic maps. That is, we write

$$\bar{z}_i(s)=g_i(z,s) \quad (5.3)$$

with the understanding that the new variables $\bar{z}(s)$ are symplectically related to the original variables $z$ for every value of the parameter $s$. Then there exists a function $h$, which we shall call the *generating function*, such that

$$\left.\frac{\partial\bar{z}_i(s)}{\partial s}\right|_z=[h(\bar{z},s),\ \bar{z}_i]. \quad (5.4)$$

*Proof*: Since the functions $g_i$ are viewed as given, we have by direct calculation

$$\frac{\partial\bar{z}_i(s)}{\partial s}=\frac{\partial g_i(z,s)}{\partial s}. \quad (5.5)$$

Next invert the transformation (5.3) to solve for the $z$'s in terms of $\bar{z}$'s, and substitute this result into the right-hand side of (5.5) to obtain expressions of the form

$$\frac{\partial\bar{z}_i(s)}{\partial s}=f_i(\bar{z},s). \quad (5.6)$$

From Taylor's theorem we have

$$\bar{z}_i(s+\epsilon)=\bar{z}_i(s)+\epsilon f_i(\bar{z},s)+O(\epsilon^2). \quad (5.7)$$

Take the Poisson bracket of (5.7) with the analogous expression having the index set equal to $j$. The result is

$$[\bar{z}_i(s+\epsilon),\bar{z}_j(s+\epsilon)]=[\bar{z}_i(s),\bar{z}_j(s)]+\epsilon\{[\bar{z}_i(s),f_j(\bar{z},s)]$$
$$+[f_i(\bar{z},s),\bar{z}_j(s)]\}+O(\epsilon). \quad (5.8)$$

Using the first part of (1.10) and equating powers of $\epsilon$, we find

$$[\bar{z}_i,f_j(\bar{z},s)]_z+[\bar{z}_j,f_i(\bar{z},s)]_z=0. \quad (5.9)$$

Here we have written the subscript $z$ to emphasize that the Poisson bracket is taken with respect to the variables $z$. However, as is well known, the Poisson bracket can also be taken with respect to the variables $\bar{z}$. For let $u$ and $v$ be any two functions. Then by the chain rule and (1.12) we have

$$[u(\bar{z}),v(\bar{z})]_z=\sum_{ij}\left(\frac{\partial u}{\partial z_i}\right)J_{ij}\left(\frac{\partial v}{\partial z_i}\right)$$
$$=\sum_{ijkl}\left(\frac{\partial u}{\partial\bar{z}_k}\right)\left(\frac{\partial\bar{z}_k}{\partial z_i}\right)J_{ij}\left(\frac{\partial v}{\partial\bar{z}_l}\right)\left(\frac{\partial\bar{z}_l}{\partial z_j}\right)$$
$$=\sum_{k,l}\left(\frac{\partial u}{\partial\bar{z}_i}\right)M_{ki}J_{ij}\tilde{M}_{jl}\left(\frac{\partial v}{\partial\bar{z}_l}\right)$$
$$=\sum_{k,l}\left(\frac{\partial u}{\partial\bar{z}_k}\right)J_{kl}\left(\frac{\partial v}{\partial\bar{z}_l}\right)=[u(\bar{z}),\ v(\bar{z})]_{\bar{z}}. \quad (5.10)$$

Thus we may also write (5.9) in the form

$$[\bar{z}_i,f_j(\bar{z},s)]_{\bar{z}}+[\bar{z}_j,f_i(\bar{z},s)]_{\bar{z}}=0. \quad (5.11)$$

The existence of the advertised generating function $h(\bar{z},s)$ now follows from Lemma 1.

*Lemma* 8: Suppose the one parameter family in Lemma 7 is also a one parameter *group*. Without loss of generality, we may assume that the parameterization is selected in such a way that it satisfies

$$\bar{z}_i(0)=z_i \quad (5.12)$$

and is additive,

$$\bar{z}_i(s_1+s_2)=g_i(\bar{z}(s_1),s_2). \quad (5.13)$$

Then the generating function $h$ is *independent* of $s$.

*Proof*: Partially differentiate (5.13) with respect to $s_2$ and then set $s_2$ equal to zero. The result is a relation of the form (5.6) with

$$f_i(\bar{z})=\left.\frac{\partial g_i}{\partial s_2}\right|_{s_2=0}. \quad (5.14)$$

Note, however, that in this case $f_i$ is independent of $s$. It follows from the remainder of Lemma 7 that the generating function $h$ is independent of $s$.

*Lemma* 9: If the generating function is independent of $s$, the differential equation

$$\frac{\partial \bar{z}_i(s)}{\partial s} = [h(\bar{z}), \bar{z}_i] \tag{5.15}$$

with the initial condition (5.12) has the unique solution

$$\bar{z}_i(s) = \exp(sH)\, z_i\,. \tag{5.16}$$

We note that apart from a sign, (5.15) is analogous to Hamilton's equations of motion (1.3).

*Proof*: Evidently (5.12) is satisfied. Now differentiate (5.16) with respect to $s$ to get

$$\frac{\partial \bar{z}_i}{\partial s} = \exp(sH) H z_i = \exp(sH)[h(z), z_i]$$
$$= [\exp(sH)\, h(z),\ \exp(sH)\, z_i]. \tag{5.17}$$

Here we have used (2.8). Also, from (2.9) and (5.16) it follows that

$$\exp(sH)\, z^\sigma = \bar{z}(s)^\sigma. \tag{5.18}$$

Consequently, since polynomials are dense in the set of functions, we must have

$$\exp(sH)\, u(z) = u(\bar{z}) \tag{5.19}$$

for any function $u$. Employing (5.16) and (5.19) in (5.17), we see that the differential equation (5.15) is satisfied.

*Lemma* 10: The function $h(z)$ is an invariant function for the transformation (5.16).

*Proof*: By (5.19) we have

$$h(\bar{z}) = \exp(sH)\, h(z) = h(z) + s[h, h] + \cdots = h(z), \tag{5.20}$$

since all the Poisson brackets are zero.

We are now ready to appreciate the result of Birkhoff, which we summarize in the next theorem.

*Theorem* 5: Denote by $\bar{z}_i(k)$ the result of applying the transformation (1.13) $k$ times in succession. We also adopt the convention (5.12). Evidently $\bar{z}_i(k)$ has an expansion similar to (1.13). We write

$$\bar{z}_i(k) = \sum_{|\sigma|>0} a_i(k, \sigma)\, z^\sigma \tag{5.21}$$

with the explicit recognition that the coefficients $a_i$ will depend on $k$. Then, in the case of two variables and provided the eigenvalues of $M(0)$ satisfy certain conditions, the dependence of the coefficients $a_i$ on $k$ can be *extended* from *integer* values to *all* real values by analytic interpolation in such a way that the series (5.21) "behaves" as a one parameter group, as in Lemma 8, with $k$ playing the role of a continuous parameter. We use the word "behaves" advisedly, because the series may not be convergent for nonintegral $k$ even though $a_i(k, \sigma)$ is well defined. That is, the group property may hold only as a formal relation between power series.

The direct verification of Birkhoff's theorem is beyond the purpose of this paper. However, we point out that his theorem, with the aid of Lemmas 9 and 10, produces a formal power series for an invariant function $h$. If the series is convergent, it yields a true invariant function; otherwise the result is a series which formally satisfies (5.20) term by term.

We will now show that the same results can be derived almost immediately for the case of any number of variables with the aid of the CBH formula. We assume the conditions of Theorem 3 are satisfied, and begin by repeatedly applying the CBH formula to the terms $\exp(G_3)\exp(G_4)\exp(G_5)\cdots$ appearing in (2.14). In view of (2.23), we can combine the exponents into one grand exponent to find

$$\exp(G_3)\exp(G_4)\cdots = \exp(G_3' + G_4' + \cdots), \tag{5.22}$$

and each term $G_r'$ can be written as a sum of a finite number of commutators. Next we try to combine the result (5.22) with $\exp(G_2)$ to find

$$\exp(G_2)\exp(G_3)\exp(G_4)\cdots = \exp(G_2)\exp(G_3' + G_4' + \cdots)$$
$$= \exp(H_2 + H_3 + H_4 + \cdots). \tag{5.23}$$

This last step is somewhat more problematical since, because of (2.23), we must now sum infinite series involving arbitrarily many commutators of $G_2$ to find each of the terms $H_3, H_4, \cdots$. We will study this matter somewhat further in a moment. Assuming that the various series converge, we can formally write

$$\bar{z}_i = \exp(H)\, z_i \tag{5.24}$$

with

$$H = H_2 + H_3 + H_4 + \cdots. \tag{5.25}$$

Furthermore, because of the relations (2.3) and (2.5), we know that there must be functions $h_2, h_3, \cdots$ corresponding to the operators $H_2$, $H_3$, etc. Thus, $H$ is a Lie operator corresponding to $h$ given by

$$h = h_2 + h_3 + h_4 + \cdots. \tag{5.26}$$

Finally, from Lemma 10 (with $s = 1$) we conclude that the function $h(z)$ constructed in this manner will be an invariant function of the transformation (2.14).

It is easy to verify that the invariant function $h$ we have obtained from the CBH formula is the same as would be found by Birkhoff's method. Let us apply the transformation (5.24) twice in succession. We write

$$\bar{z}_i = \exp(H_z)\, z_i, \tag{5.27a}$$

$$\bar{\bar{z}}_i = \exp(H_{\bar{z}})\, \bar{z}_i, \tag{5.27b}$$

and use subscripts to indicate exactly which variables occur in the various Poisson brackets. Expanding (5.27b) we can write

$$\bar{\bar{z}}_i = \exp(H_{\bar{z}})\, \bar{z}_i = \bar{z}_i + [h(\bar{z}), \bar{z}_i]_{\bar{z}} + \cdots$$
$$= \bar{z}_i + [h(\bar{z}), \bar{z}_i]_z + \cdots = \bar{z}_i + [h(z), \bar{z}_i]_z + \cdots$$
$$= \exp(H_z)\, \bar{z}_i\,. \tag{5.28}$$

Here we have used (5.10) and the fact that $h$ is an invariant function. Now substitute (5.28) into (5.27a) to get

$$\bar{\bar{z}}_i = \exp(H_z)\exp(H_z)\, z_i = \exp(2H_z)\, z_i\,. \tag{5.29}$$

Employing the notation of Theorem 5, it is clear that (5.29) generalizes to

$$\bar{z}_i(k) = \exp(kH)\, z_i. \tag{5.30}$$

Finally, (5.30) can be extended from integer values to all real values simply by replacing $k$ with $s$ to give (5.16).

There is an interpretation of the result (5.30) which is worth emphasizing: We have already remarked that (5.15) is analogous to Hamilton's equations of motion. We now see that if Birkhoff's theorem, or equivalently the use of the CBH formula, is applicable, then the transformation (1.13) can be viewed as the result of integrating Hamilton's equations of motion for the time independent Hamiltonian $(-h)$ from the initial "time" $s=0$ to the "time" $s=1$. Subsequent iterations of the map are obtained by integrating on to successive integer values.

So far, we have not discussed the convergence of the various procedures we have employed. This question is very difficult, and much remains to be learned. A theorem of Moser[27] can be used in the simplest case of two variables $z_1, z_2$ if $M(0)$ can be brought to the form (3.3). In our language, he shows in this case by indirect methods that if $\mathcal{T}$ is the symplectic transformation in question, then there exists another symplectic transformation $\mathcal{U}$ of the form

$$\mathcal{U} = \exp(F_3)\exp(F_4)\exp(F_5)\cdots \tag{5.31}$$

such that

$$\mathcal{U}^{-1}\mathcal{T}\mathcal{U} = \exp\left(\sum_1^\infty \alpha_n G_{2n}\right) \tag{5.32}$$

with

$$g_2 = z_1 z_2, \quad g_{2n} = (g_2)^n, \quad \alpha_1 = -\log\lambda. \tag{5.33}$$

Both the infinite product (5.31) and the infinite series in (5.32) converge. By undoing the transformation $\mathcal{U}$, one finds the desired result

$$\mathcal{T} = \exp H \tag{5.34}$$

with

$$H = \mathcal{U}\sum_1^\infty \alpha_n G_{2n}\,\mathcal{U}^{-1}. \tag{5.35}$$

Thus, there are nontrivial classes of problems for which our methods (and also Birkhoff's) succeed.

By contrast, a theorem of Moser's on Cremona maps[22] can be used to infer that the CBH series diverges for the example (3.2). The method of proof is again indirect. However, direct examination of the CBH series shows that it repeatedly contains terms of the form

$$[1-\exp(-\hat{G}_2)]^{-1}F_n, \tag{5.36}$$

and we will see that these terms can cause problems.

Rather than examining (5.36), it is convenient to use the homomorphism between Lie algebras and their adjoints to work instead with the expression

$$[1-\exp(-G_2)]^{-1}f_n. \tag{5.37}$$

We next observe from (2.23) that $G_2$ maps $\mathcal{P}_n$ into itself, and hence the action of $G_2$ on each $\mathcal{P}_n$ can be represented by a matrix. Let $v_1$ be a polynomial of first degree which is an eigenvector of $G_2$. We write the eigenvalue as $(-\log\lambda)$ so that we have

$$G_2 v_1 = (-\log\lambda)\, v_1 \tag{5.38}$$

and

$$\exp(-G_2)\, v_1 = \lambda v_1. \tag{5.39}$$

Now suppose $f_n$ is a polynomial of the form

$$f_n = (v_1)^n. \tag{5.40}$$

Then we find

$$\exp(-G_2) f_n = \exp(-G_2)(v_1)^n = \lambda^n (v_1)^n = \lambda^n f_n. \tag{5.41}$$

Here we have used (2.7). In this case we find for (5.37) the result

$$[1-\exp(-G_2)]^{-1} f_n = (1-\lambda^n)^{-1} f_n. \tag{5.42}$$

Suppose $\lambda$ lies on the unit circle as it does for the example (3.2). Then the expression $(1-\lambda^n)^{-1}$ either is infinite for some $n$ [if $\alpha = i\log\lambda$ is a rational multiple of $2\pi$], or becomes arbitrarily large with increasing $n$ [if $\alpha$ is an irrational multiple of $2\pi$]. What we are observing here is a manifestation of the classic problem of "small denominators" which has been known to celestial mechanicians in connection both with perturbation theory and mapping problems since the time of Poincaré.[28,29] We see that it may also occur in the Lie algebraic approach in such a way as to spoil the convergence of the CBH series, and that this problem can potentially occur if any of the eigenvalues of $M(0)$ lie on the unit circle.

There is one last topic we wish to discuss. We have used the CBH series to obtain an invariant function $h$. In the case of sympletic transformations in two variables, a single invariant function suffices to characterize the map, and all other invariant functions are simple functions of $h$. However, in the case of four or more variables, e.g., (3.27), there may be additional invariant functions beyond $h$.

In view of (5.19), $f$ will be an invariant function if it satisfies the relation

$$Hf = [h, f] = 0. \tag{5.43}$$

Consequently, the problem of finding further invariant functions is equivalent to the classical mechanics problem of finding integrals of motion for a system having $(-h)$ as a Hamiltonian. By analogy to classical mechanics, we expect to be able to find at most $2n-1$ functionally independent invariant functions including $h$ itself.

There is as yet no fully developed algorithm for finding integrals of motion for any specified Hamiltonian $h$. However, there is a germ for such an algorithm in Birkhoff's procedure of attempting to bring Hamiltonians to a normal form.[4,5] In our notation, one attempts in this procedure to find polynomials $g_3$, $g_4$, etc., such that the Hamiltonian $h'$ given by

$$h' = \cdots \exp(G_4)\exp(G_3)\, h \tag{5.44}$$

has a particularly simple form. If the form is simple enough, one can read off the integrals of motion directly. This method has been applied successfully by Gustavson and others[5,6] to the case of several variables provided $h_2$ has the form

$$h_2 = \sum_i \alpha_i z_i^2 \tag{5.45}$$

with all $\alpha_i > 0$. The case with some $\alpha_i = 0$ can also be treated.[17] An analysis which we intend to publish later shows that what is essential to this whole procedure is a detailed treatment of the range and null spaces of the operator $H_2$.

Now suppose that $f'$ is an integral of $h'$,

$$[h', f'] = 0. \tag{5.46}$$

We define $f$ by the rule

$$f = \exp(-G_3)\exp(-G_4)\cdots f'. \tag{5.47}$$

Then we find

$$\begin{aligned}[h,f] &= [\exp(-G_3)\exp(-G_4)\cdots h', \exp(-G_3)\exp(-G_4)\cdots f'] \\ &= \exp(-G_3)\exp(-G_4)\cdots [h', f'] = 0,\end{aligned} \tag{5.48}$$

which shows that $f$ is an integral. Usually $f'$ can be taken to be a second degree polynomial. However the series $f$ given by (5.47) will generally contain an infinite number of terms and may not converge. In the latter case, (5.43) is only satisfied term by term, and $f$ is only a formal series. We expect the case of divergent series to be the most common. This is because if the series were to converge, it would produce an analytic global integral for the Hamiltonian $h$. However, most Hamiltonians do not possess global analytic integrals.[7,29]

## 6. CONCLUDING SUMMARY

In Sec. 2 it was shown that the Lie transformation associated with an analytic function produces an analytic symplectic map, and that conversely, under certain general conditions, an analytic symplectic map can be written as a product of Lie transformations. Section 3 treated several examples of analytic symplectic maps that had been studied previously by other authors. The discussion then turned in Sec. 4 to a further development of Lie algebraic tools and culminated with the Campbell—Baker—Haudsdorff formula. Next, after some preliminary background work, it was shown in Sec. 5 that the CBH formula can be used to formally combine a product of Lie transformations into a single Lie transformation, and that in so doing one obtains a generalization of Birkhoff's theorem for the construction of invariant functions. Thus, the existence of invariant functions is intimately related to the convergence of the CBH formula, and vice versa. Finally, in the case of symplectic maps involving more than two variables, the construction of additional invariant functions was shown to be analogous to the construction of integrals of motion in Hamiltonian dynamics.

## ACKNOWLEDGMENTS

The first author is indebted to the Institut des Hautes Etudes Scientifiques where this work was begun. He wishes to thank Professor N. Kuiper and Professor L. Michel for their fine hospitality.

## APPENDIX A

The purpose of this appendix is to demonstrate that $M$ as given by (3.37) lies on a one parameter subgroup connected to the identity. We begin by observing that $M$ can be written as a product of two symplectic matrices $N$ and $R$,

$$M = NR, \tag{A1}$$

where

$$N = \begin{pmatrix} 1 & 0 & a_1+b & b \\ 0 & 1 & b & a_2+b \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{A2}$$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}. \tag{A3}$$

Each of these matrices can be written in exponential form, and use of Lemma 3 reveals that they lie on one parameter symplectic subgroups continuously connected to the identity,

$$N = \exp\begin{pmatrix} 0 & Q \\ 0 & 0 \end{pmatrix}, \tag{A4}$$

$$R = \exp\begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix}, \tag{A5}$$

Here each block is a $2\times 2$ matrix, and $Q$ denotes the matrix

$$Q = \begin{pmatrix} a_1+b & b \\ b & a_2+b \end{pmatrix}. \tag{A6}$$

The next step is to try to combine the two exponents by using the CBH formula. For this purpose we note that the two exponents occurring in (A4) and (A5) can be written in the respective forms

$$\log N = Q \otimes (\sigma_1 + i\sigma_2)/2 \tag{A7}$$

$$\log R = I \otimes (\sigma_1 - i\sigma_2)/2. \tag{A8}$$

Here the symbols $\sigma_j$ denote the Pauli matrices,[1] and "$\otimes$" indicates that we have taken a tensor product.[30] For example,

$$Q \otimes (\sigma_1 + i\sigma_2)/2 = Q \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & Q \\ 0 & 0 \end{pmatrix}. \tag{A9}$$

It is easily verified that the tensor product operation obeys the multiplication rules

$$(A \otimes \sigma_j)(B \otimes \sigma_k) = AB \otimes \sigma_j \sigma_k \tag{A10}$$

and the addition rules

$$(A+B) \otimes \sigma_j = A \otimes \sigma_j + B \otimes \sigma_j, \tag{A11}$$

$$A \otimes (\alpha\sigma_j + \beta\sigma_k) = \alpha A \otimes \sigma_j + \beta A \otimes \sigma_k. \tag{A12}$$

Consequently, from the CBH formula we conclude that $\log M$ must be given by an expression of the form

$$\log M = \log NR = f \otimes \sigma_1 + g \otimes \sigma_2 + h \otimes \sigma_3, \tag{A13}$$

where $f$, $g$, and $h$ are power series in the matrix $Q$. This is because $I$ and powers of $Q$ all commute, and the Pauli matrices are closed under commutation. Furthermore $f$, $g$, and $h$ must be the same series that occur in the expression

$$\exp[\alpha(\sigma_1 + i\sigma_2)/2]\exp[(\sigma_1 - i\sigma_2)/2] = \exp[f(\alpha)\sigma_1 + g(\alpha)\sigma_2 + h(\alpha)\sigma_3], \qquad \text{(A14)}$$

where $\alpha$ is a parameter.

It remains to be shown that the series $f$, $g$, and $h$ converge for $Q$ sufficiently small. Or equivalently, we must show that $f(\alpha)$, $g(\alpha)$, and $h(\alpha)$ are all analytic and have nonzero radii of convergence in the complex variable $\alpha$. A short calculation for the group SL(2, C) gives the multiplication rule

$$\exp(\mathbf{n}_1 \cdot \boldsymbol{\sigma})\exp(\mathbf{n}_2 \cdot \boldsymbol{\sigma}) = \exp(\mathbf{n}_3 \cdot \boldsymbol{\sigma}), \qquad \text{(A15)}$$

with $\mathbf{n}_3$ given by the formulas

$$\boldsymbol{\tau}_3 = (\boldsymbol{\tau}_1 + \boldsymbol{\tau}_2 + i\boldsymbol{\tau}_1 \times \boldsymbol{\tau}_2)/(1 + \boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2), \qquad \text{(A16)}$$

$$\boldsymbol{\tau}_j = \mathbf{n}_j(\tanh\sqrt{\mathbf{n}_j \cdot \mathbf{n}_j}/\sqrt{\mathbf{n}_j \cdot \mathbf{n}_j}) = \mathbf{n}_j(1 - \tfrac{1}{3}\mathbf{n}_j \cdot \mathbf{n}_j + \cdots). \qquad \text{(A17)}$$

For the case in question, Eq. (A14), we have

$$\mathbf{n}_1 = \alpha(\hat{e}_1 + i\hat{e}_2)/2, \quad \mathbf{n}_2 = (\hat{e}_1 - i\hat{e}_2)/2. \qquad \text{(A18)}$$

Inserting this information into (A16) and (A17) gives the result

$$\boldsymbol{\tau}_3 = [\hat{e}_1(1+\alpha)/2 + i\hat{e}_2(\alpha - 1)/2 + \hat{e}_3\alpha/2]/[1 + \alpha/2]. \qquad \text{(A19)}$$

Now we need to solve for $\mathbf{n}_3$. Combining (A19) and (A17) we find

$$\boldsymbol{\tau}_3 \cdot \boldsymbol{\tau}_3 = \alpha(1 + \alpha/4)/(1+\alpha/2)^2 = (\tanh\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3})^2 = \mathbf{n}_3 \cdot \mathbf{n}_3 - \tfrac{2}{3}(\mathbf{n}_3 \cdot \mathbf{n}_3)^2 + \cdots. \qquad \text{(A20)}$$

Consequently,

$$(\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3}/\tanh\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3}) = 1 + (\alpha/3) + \cdots. \qquad \text{(A21)}$$

Finally, from the relation

$$\mathbf{n}_3 = \boldsymbol{\tau}_3(\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3}/\tanh\sqrt{\mathbf{n}_3 \cdot \mathbf{n}_3}) \qquad \text{(A22)}$$

we find

$$f(\alpha) = (1/2) + (5\alpha/12) + \cdots, \qquad \text{(A23)}$$
$$g(\alpha) = (-i/2) + (7i\alpha/12) + \cdots,$$
$$h(\alpha) = (\alpha/2) - (\alpha^2/12) + \cdots.$$

It is clear from (A19)–(A22) that the series for $f$, $g$, and $h$ all have nonzero radii of convergence.

## APPENDIX B

The purpose of this appendix is to prove Theorem 4. Writing (4.22) a bit more explicitly, we have

$$\exp[C(\alpha,\beta)] = \exp(\alpha A)\exp(\beta B). \qquad \text{(B1)}$$

The first result we will need is that $C$ obeys the differential equation

$$\frac{\partial C}{\partial \beta} = \hat{C}[1 - \exp(-\hat{C})]^{-1}B. \qquad \text{(B2)}$$

To see how this comes about, let us differentiate both sides of (B1) with respect to $\beta$. The derivative of the right-hand side is easily computed,

$$\frac{\partial}{\partial\beta}\exp(\alpha A)\exp(\beta B) = \exp(\alpha A)\exp(\beta B)B = \exp(C)B. \qquad \text{(B3)}$$

Computation of the derivative of the left-hand side requires more work. We find through first order in $\delta\beta$ that

$$\exp[C(\alpha, \beta + \delta\beta)] = \exp\left(C(\alpha,\beta) + \delta\beta\frac{\partial}{\partial\beta}C(\alpha,\beta)\right) = \sum_0^\infty (1/n!)\left(C(\alpha,\beta) + \delta\beta\frac{\partial}{\partial\beta}C(\alpha,\beta)\right)^n. \qquad \text{(B4)}$$

Now expand the power series and retain zero and first order terms. The result through first order is

$$\exp[C(\alpha,\beta+\delta\beta)] = \exp(C) + \delta\beta\sum_{n=1}^\infty\sum_{m=0}^{n-1}(1/n!)C^m\left(\frac{\partial C}{\partial\beta}\right)C^{n-m-1}. \qquad \text{(B5)}$$

Here we have paid careful attention to the possibility that $C$ and $\partial C/\partial\beta$ may not commute. From (B5) we conclude

$$\left(\frac{\partial}{\partial\beta}\right)\exp(C) = \sum_{n=1}^\infty\sum_{m=0}^{n-1}\frac{1}{n!}C^m\left(\frac{\partial C}{\partial\beta}\right)C^{n-m-1}. \qquad \text{(B6)}$$

Next change the order of summation in (B6) to obtain

$$\left(\frac{\partial}{\partial\beta}\right)\exp(C) = \sum_{m=0}^\infty\sum_{l=0}^\infty[1/(l+m+1)!]C^m\left(\frac{\partial C}{\partial\beta}\right)C^l. \qquad \text{(B7)}$$

It is a remarkable fact that the series (B7) has an integral representation,

$$\sum_{m=0}^\infty\sum_{l=0}^\infty[1/(l+m+1)!]C^m\left(\frac{\partial C}{\partial\beta}\right)C^l = \int_0^1 d\gamma\exp[(1-\gamma)C]\left(\frac{\partial C}{\partial\beta}\right)\exp(\gamma C). \qquad \text{(B8)}$$

This is easily verified by expanding out the two exponentials and integrating term by term.

Only a few more steps are required. We write

$$\int_0^1 d\gamma\exp[(1-\gamma)C]\left(\frac{\partial C}{\partial\beta}\right)\exp(\gamma C) = \exp(C)\int_0^1 d\gamma\exp(-\gamma C)\left(\frac{\partial C}{\partial\beta}\right)\exp(\gamma C) = \exp(C)\int_0^1 d\gamma\exp(-\gamma\hat{C})\left(\frac{\partial C}{\partial\beta}\right) = \exp(C)\{[1-\exp(-\hat{C})]/\hat{C}\}\left(\frac{\partial C}{\partial\beta}\right). \qquad \text{(B9)}$$

Here we have used (4.6). Also, the last integration over $\gamma$ was performed by expanding $\exp(-\gamma\hat{C})$ in a power series, integrating term by term, and then resuming the result. We conclude that

$$\left(\frac{\partial}{\partial\beta}\right)\exp(C) = \exp(C)\{[1-\exp(-\hat{C})]/\hat{C}\}\left(\frac{\partial C}{\partial\beta}\right). \qquad \text{(B10)}$$

We note for future use that Eqs. (B4)–(B10) hold quite generally, and make no use of any special form $C$ might have.

The differentiation of both sides has been completed. Comparing (B3) and (B10) and cancelling the common factor $\exp(C)$, we find

$$\{[1-\exp(-\hat{C})]/\hat{C}\}\ \frac{\partial C}{\partial \beta} = B. \tag{B11}$$

This expression, when solved for $\partial C/\partial \beta$, gives the advertised result (B2).

The proof of Eq. (4.22) now follows immediately. Make the expansion

$$C(\alpha,\beta)=\sum \alpha^m \beta^n C_{mn} \tag{B12}$$

and substitute it into (B2) with the observation that

$$C(\alpha, 0)=\alpha C_{10}=\alpha A, \tag{B13}$$

$$C(0,\beta)=\beta C_{01}=\beta B. \tag{B14}$$

A comparison of coefficients of like powers of $\alpha^m\beta^n$ gives the series (4.23).

To prove Eq. (4.24) we write a Taylor expansion of $C$ with respect to $\beta$,

$$C(\beta)=C(0)+\beta C'(0)+(\beta^2/2)C''(0)+\cdots. \tag{B15}$$

Here we have suppressed the dependence of $C$ on $\alpha$ for notational convenience. The quantity $C(0)$ is already known from (B13), and $C'(0)$ can be found from (B2) with $\beta$ set equal to zero. The result is

$$C'(0)=\alpha\hat{A}[1-\exp(-\alpha\hat{A})]^{-1}B. \tag{B16}$$

Insertion of (B13) and (B16) into (B15) gives the desired result (4.24).

The computation of successively higher derivatives becomes increasingly more complicated. To find $C''(0)$, we write equation (B11) in the form

$$\mathcal{P}(\beta)\,C'(\beta)=B, \tag{B17}$$

where $\mathcal{P}(\beta)$ denotes the operator

$$\mathcal{P}(\beta)=[1-\exp(-\hat{C})]/\hat{C}=\int_0^1 d\nu \exp(-\nu\hat{C}). \tag{B18}$$

Next we differentiate both sides of (B17) with respect to $\beta$ to find

$$\mathcal{P}'C'+\mathcal{P}C''=0, \tag{B19}$$

and hence

$$C''(0)=-\mathcal{P}^{-1}(0)\,\mathcal{P}'(0)\,C'(0). \tag{B20}$$

From the integral representation (B18) we find

$$\mathcal{P}'(\beta)=-\int_0^1 d\nu \exp(-\nu\hat{C})\int_0^1 d\gamma\exp(\gamma\nu\hat{C}) \times\nu\hat{C}'(\beta)\exp(-\gamma\nu\hat{C}). \tag{B21}$$

Here we have also used (B4)—(B8) and part of (B9) to differentiate $\exp(-\nu\hat{C})$. It follows that

$$-\mathcal{P}'(0)\,C'(0)=\int_0^1\int_0^1 \nu\, d\nu\, d\gamma \exp(-\nu\alpha\hat{A})\exp(\gamma\nu\alpha\hat{A}) \times\hat{C}'(0)\exp(-\gamma\nu\alpha\hat{A})\,C'(0). \tag{B22}$$

We also observe that

$$\begin{aligned}\hat{C}'(0)\exp(-\gamma\nu\alpha\hat{A})\,C'(0) &=\{C'(0),\exp(-\gamma\nu\alpha\hat{A})\,C'(0)\}\\ &=\exp(-\gamma\nu\alpha\hat{A})\{\exp(\gamma\nu\alpha\hat{A})\,C'(0),C'(0)\}.\end{aligned} \tag{B23}$$

In obtaining the last expression we have used a result analogous to (2.8). By combining (B18), (B20), (B22), and (B23) we find the final result

$$C''(0)=\alpha\hat{A}[1-\exp(-\alpha\hat{A})]^{-1}\int_0^1\int_0^1 \nu\, d\nu\, d\gamma\exp(-\nu\alpha\hat{A}) \times\{\exp(\gamma\nu\alpha\hat{A})\,C'(0),C'(0)\}. \tag{B24}$$

*Supported in part by the National Science Foundation under Grant No. GP-41822X.

†Present address: Laboratory of Plasma Studies, Cornell University, Ithaca, New York 14853.

[1]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading Massachusetts, 1950).
[2]E. Saletan and A. Cromer, *Theoretical Mechanics* (Wiley, New York, 1971).
[3]E.C.G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective* (Wiley, New York, 1974).
[4]G.D. Birkhoff, *Dynamical Systems*, Am. Math. Soc. No. 9 (Am. Math. Soc., New York, 1927).
[5]A. Deprit, J. Henrard, J. Price, and A. Rom, Celest. Mech. **1**, 222 (1969).
[6]F. Gustavson, Astron. J. **71**, 670—86 (1966).
[7]A.J. Dragt and J.M. Finn, J. Geophys. Res. **81**, 2327—40 (1976).
[8]V.I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).
[9]J.M. Finn, Nucl. Fusion **15**, 845 (1975); N. Filonenko, R. Sagdeev, G. Zaslavski, Nucl. Fusion **7**, 253 (1967); M.N. Rosenbluth, R.Z. Sagdeev, J.B. Taylor, and G.M. Zaslavski, Nucl. Fusion **6**, 297 (1966); D.W. Kerst, J. Nucl. Energy Part C **4**, 253—62 (1962).
[10]A. Deprit, Celest. Mech. **1**, 12 (1969).
[11]A.H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973) E. Leimanis, *The General Problem of the Motion of Coupled Rigid Bodies About a Fixed Point* (Springer, New York, 1965), p. 121.
[12]G.E.O. Giacaglia, *Perturbation Methods in Non-Linear Systems* (Springer-Verlag, Berlin, 1972).
[13]R. Abraham, *Foundations of Mechanics* (Benjamin, New York, 1967).
[14]M. Hausner and J.T. Schwartz, *Lie Groups; Lie Algebras* (Gordon and Breach, New York, 1968).
[15]S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[16]W. Engel, Math. Ann. **130**, 11 (1955); **136**, 319 (1958).
[17]J.M. Finn, University of Maryland Thesis (unpublished 1974).
[18]C. Froeschle, Astron. Astrophys. **9**, 15 (1970).
[19]M. Henon, Q. Appl. Math. **27**, 291 (1969); J. Roels and M. Henon, Bull. Astron. Ser. 3 **2**, 267 (1967); M. Henon and C. Heiles, Astron. J. **69**, 73 (1964).
[20]J.H. Bartlett, *Classical and Modern Mechanics* (University of Alabama Press, Birmingham, Alabama, 1975); B.Z. Jenkins and J.H. Bartlett, Celest. Mech. **5**, 407 (1972).
[21]J. Greene, J. Math. Phys. **9**, 760 (1968).
[22]J. Moser, Bol. Soc. Mat. Mex. **5**, 176 (1960).
[23]F. Rannou, Astron. Astrophys. **31**, 289 (1974).
[24]C. Froeschle, Astron. Astrophys. **9**, 15 (1970); **16**, 172 (1972); Astrophys. Space Sci. **14**, 110 (1971); C. Froeschle and J.P. Scheidecker, Astron. Astrophys. **22**, 431 (1973).
[25]N. Bourbaki, *Elements de mathêmatiques, Fasc.* 37: *Groupes et algebres de Lie* (Hermann, Paris, 1972), Chapters 2 and 3.
[26]G.D. Birkhoff, Acta Math. **43**, 1 (1920).
[27]J. Moser, Commun. Pure Appl. Math. **9**, 673 (1956).
[28]H. Poincaré, *Les methods nouvelles de la mechanique celeste*, 1892 (Reprinted by Dover, New York, 1957).
[29]J. Moser, *Stable and Random Motions in Dynamical Systems* (Princeton U.P., Princeton, N.J., 1973).
[30]P. Halmos, *Finite Dimensional Vector Spaces* (Van Nostrand, Princeton, N.J., 1958).